MINAR International Journal of Applied Sciences and Technology
ISSN: 2717-8234
Article type: Research Article

Received: 12/10/2022 **Accepted:** 01/11/2022 **Published:** 01/12/2022

DATA MODELLING OF LC/MS-BASED METABOLOMIC PROFILING TO COMPARE BETWEEN HUMAN PLASMA AND URINE SAMPLES ASSOCIATED WITH BEETROOT JUICE

Ali Muhsen ALI [1]

University of Kerbala, Iraq

Amel D. HUSSEIN[2]

Wasit University, Iraq

Abstract:

Data modelling- based untargeted metabolomic researches is one of the best approaches which can be used to compare among biological fluid samples to provide a comprehensive and reliable sight about the changes of metabolomic profiling. This study sought to compare between human urine and plasma to investigate metabolomic changes of a diet pre- and post-intake beetroot juice that offer unique metabolomic fingerprint associated to the potential effects of beetroot juice. A current pilot study of metabolomic patterns used Liquid Chromatography coupled to Mass Spectrometer (LC-MS) to carry out an analysis for seventy-two plasma and urine samples, equally. Samples were collected from nine adult healthy subjects (4 samples per subject) at pre-, as baseline, and post-intake beetroot juice in three stages, after 2hrs, 4hrs, and 8hrs. On the basis of the validation of data modelling, robust separation was observed between urine samples pre and post-intake beetroot juice and was more fitting and significant than the separation between plasma samples. The results of pilot study indicate that metabolomics screening of urine samples may be the best tool and a potential approach to predict the metabolomic profiling than plasma samples to assess the metabolic effect of a diet pre- and post-intake beetroot juice. As a result of the effects of beetroot juice, the present results also uncover significantly changes in most important metabolites including amino acid, peptide, Co-factors and vitamins which may contribute to the consolidation of the using of plant metabolites and natural substances to synthesis the nanoparticles for the biomedical applications.

Keywords: Metabolomic Profiling, Beetroot Juice (BJ), Biological Fluid, LC-MS, Principal Components Analysis (PCA), Orthogonal Partial Least Squares- Discriminant Analysis (OPLS-DA)..

http://dx.doi.org/10.47832/2717-8234.13.8

[1] ali.muhsen@uokerbala.edu.iq, https://orcid.org/0000-0001-7275-4942

[2] adashar@uowasit.edu.iq

This article has been scanned by **iThenticat** No **plagiarism** detected

Introduction:

Nutrition studies reported that high nitrate vegetables have special and different role in determination of relation between biological systems and nutritional status **[1]**. Beetroot as well as many types of the vegetables such as celery, lettuce, spinach and rocket…etc., have been classified as nitrate-rich vegetables which can rises the levels of circulating metabolites of nitric oxide (NO). Beetroot is one of the common popular vegetable in wide rang about of world as a part of healthy diet packed with nutrition and plays an important role as antioxidants which help in prevent cell damage and reduce the risk of many disease. It can be consumed raw or used in other different forms such as its using in the food industry for the production of dry and frozen preserves, as concentrated fruit juices, or as food additives **[2]**. Many previous studies have suggested that foods with the highest nitrate contents may contribute in fighting of many of chronic disease such as coronary heart disease **[3]**, ischemic stroke **[4]**, high blood pressure **[5]** and diabetes mellitus **[6]**. Also, a number of specific biomarkers for foods such as meat and vegetables that can be used as indicators in prediction the negative or positive influence for these types of foods were exposed by other nutrition studies **[7]**.

However, the effect of nutritional factors on human biological systems which reflex the complex relationship between nutrients and metabolism processes remains to be understood. Recently, nutrition-based metabolomics research seek well-established the true association between dietary intake and its influence human and animal health through assessment the metabolic effects of dietary. This assessment provide the tries to increase understanding of metabolic pathways changes during diet involved in this relationship between nutrients and disease progression. Recent applications of metabolomics studies of nutrition, metabolism and lipid dysfunction were covered by M. Oresic in his review **[8]**. According to develop the uses of metabolomic strategy in this field, other authors in 2016 reviewed many of nutrition studies and listed them base on metabolomic technique, dietary factor, study duration, samples and biomarker **[9]**.

Based on the hypothesis tested and goal of metabolome study, metabolomic analytical strategy has been divided to two strategies included a targeted and untargeted approach. Also the type and easy availability of sample are main factors which be taken into consideration to study the changes of metabolomic profiling by these strategies. Thus, choosing of the potential specimen and appropriate approach become essential key in the application of metabolomic analysis to be the better tool to enhance understanding of the relation between nutrition and human metabolomic profiling. For this reason, this study was designed to compare between two common human biological fluids (urine and plasma samples) to choose the best tools which can be used to assess the metabolic effect of a diet associated with beetroot juice.

Materials and Methods:

Chemicals and Solvents:

HPLC grade Acetonitrile (ACN) and formic acid (98%) were purchased from Fisher Scientific (Loughborough, UK) and BDH-Merck (Poole, UK), respectively. A Direct-Q 3 Ultrapure Water System (Millipore, Watford, UK) was used to produce HPLC grade water. Ammonium carbonate and methanol (MeOH) were obtained from Sigma-Aldrich (Poole, UK).

Subjects and Experimental Design:

Seventy-two human plasma and urine samples, equally, were collected from nine healthy and non-smoker adult volunteers (age 30 ± 5 years, stature: 180 ± 8 cm, body mass: 83.4 ± 10.4 kg). After checking of family history for all volunteers, they were underwent many

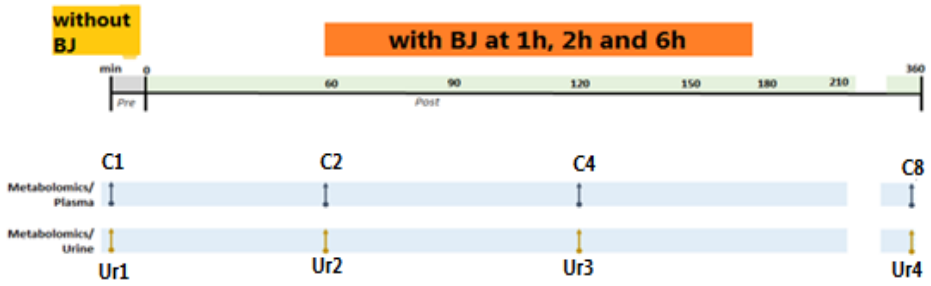
investigations to exclude any person had a current illness or viral infection within one month ago, and had a known disorder as hypertensive (140/90 mmHg) or premature cardiovascular disease.

All participants in the study provided written informed consent and a medical questionnaire before the study began. Specimen of subjects (plasma and urine) were taken as four samples for each specimen per subject in same day and were divided to one specimen pre-, as baseline, and three specimens post-intake beetroot juice in three stages, after 1hr, 2hrs, and 6hrs. Collection of samples and schematic of study are summarised in figure 1. Plasma sample was immediately separated from centrifuged blood after it's collected at 5000rpm for 10min and stored at -20°C. Urine samples were collected in urine containers, it labelled with participants name, and delivered to the laboratory for storage at -20°C until the day of analysis.

Plasma & Urine Samples preparation:

Preparation of sample in metabolomics analysis mainly depends on type of samples and analysis methods **[10]**. Plasma and urine samples were stored at -20°C and thawed at room temperature before its preparation. On analysis day, these samples were prepared for LC-MS analysis by unfreezing them at room temperature. For urine samples preparation, 200ul of each urine were thoroughly mixed with 800µl of CAN to precipitate urine protein content. Then, the solution was thoroughly mixed by a vortex machine and centrifuged for 10 -15min with 15000 rpm at 4°C.

The clear solution, supernatant, from each samples was transferred to the relevant HPLC vials to be ready for LC-MS analysis. In the identification of metabolites plasma samples, the protein precipitation by methanol may be the most appropriate way to treatment the protein content in plasma samples in LC-MS approaches to avoid damage to the analytical column and MS capillaries **[11]**. Therefore, the plasma preparation was performed by dilution of 200ul of each plasma sample with 800ul of MeOH/CAN (80/20) followed by shaking and centrifugation at 4°C for 10 -15min/15000 rpm. Then, 800µl of supernatant for each sample was transferred onto the correspondingly labelled HPLC vials for analysis.



Specimen	Time: Baseline (pre)	60mins (post-1hr)	120mins (post-2hrs)	360mins (post-6hrs)
Plasma	C1	C2	C4	C8
Urine	Ur1	Ur2	Ur3	Ur4

Figure 1: Indicative representation of plasma and urine specimens' collection schematic at four stages for two conditions. Firstly without dietary conditions as baseline stage (pre-intake beetroot juice) and secondly with dietary conditions (post-intake beetroot juice) in three stages at 1hr, 2hrs, and 6hrs

HILIC–HRMS Analysis Conditions:

LC-MS-based plasma and urine samples analysis was performed on an Dionex 3000 HPLC (Thermo Fisher Scientific, Hemel Hempstead, UK) combined with an Exactive Orbitrap (Thermo Fisher Scientific) in both positive and negative mode set at 50,000 resolution (controlled by Xcalibur version 2.1.0; Thermo Fisher Scientific, Hemel Hempstead, UK). The mass range (m/z) was scanned at 75–1200 and the capillary temperature was 320°C as well as the flow rates of auxiliary gas and sheath were 17 and 50 arbitrary units, respectively. The separation was carried out by injection 10µl of each sample solution on a zwitterionic-hydrophilic interaction chromatography column, ZIC-pHILIC column, (150 mm x 4.6 mm; 5µm from HiChrom, Reading, UK) with mobile phase of (A): 20 mM ammonium carbonate in HPLC grade water (pH 9.2), and (B): HPLC grade acetonitrile (CAN).

Finally, The prepared urine and plasma sample solutions were kept in a vial tray which was set at 4°C and a flow rate of mobile phase was 300 µL/min in binary gradient mode which was as follows: 80% of B at 0 min, 20% B at 30 min, 20% B at 36 min, 80% B at 37 min and 80% B 46 min **[12 & 13]**.

LC-MS Data Processing and Statistical Analysis:

M/Z mine 2.14 was used to extract raw LC-MS files obtained from Xcalibur software **[14]** in order to metabolite identification by peak extraction and alignment, as previously described. Data from the Human Metabolome Database (HMD), Lipid Maps and the Metlin Database were collected to prepare unified database called House Metabolite Database which used for search the accurate masses and detect a putative identification of metabolites **[15]**. Background peaks in the blank in MZmine were removed before data transferring for carrying out univariate analysis. Univariate analysis including normalisation, the area for each metabolite divided to the mean of the peak areas, for each metabolite, across the samples with and without beetroot juice. Statistical analysis, paired t-test (p-value) and fold changes (ratio), were completed in Excel (Microsoft Office 2013). SIMCA-P version 14.0 (Umetrics, Sweden) was applied to conduct the multivariate analysis for data modelling to build the models of principal components analysis (PCA), an unsupervised analysis method, and Orthogonal Partial least squares- discriminant analysis (OPLS-DA), a supervised method, **[16]**. For determining of significant influence of the components in the dataset, S-plots was generated by centring of data and extraction of Pareto scaled for both models, PCA and OPLS-DA.

Results and Discussion:

The first and important step in study of metabolomic changes is correct selecting of biological samples to achieve wanted goal through understanding these complex changes in metabolic pathways.

Pathway and metabolites concentration changes:

Experimental studies of metabolic patterns play a major role that researchers use to select appropriate samples to determine changes of metabolic. The LC-MS method was used to identify the plasma and urine metabolites which change before and after intake beetroot juice by healthy humans.

In plasma and urine samples from this study, hundreds of metabolites with minimal masses and retention times (deviation <3 ppm, MSI levels 1 or 2) were identified by searching the database and matching standards against MZmine 2.14. For comparison among metabolites on the same axis to obtain the best modeling for data comparison, the peak areas

for each metabolite at all time points were normalized as it was applied and confirmed in our previous study **[15]**.

Paired t-tests and fold changes (ratios) were performed as univariate analysis to compare among four time samples at pre-, and during 2hrs, 4hrs, and 8hrs post-intake beetroot juice to determine that differences were intrinsically related to the samples taken. In univariate analysis, a one-way analysis of the original data was processed by calculating the above ratio. This comparison between samples indicated that metabolic changes evident occurred between the time points of baseline (pre-intake) and treated-line (post-intake) to be goal of building a multivariate model of current study approach.

Thus, the identification of metabolites and their pathways in urine and plasma samples which altered in response to nutrition with beetroot juice were highlighted. Data filtering by normalization and one-way analysis revealed significant changes in concentration of metabolites in urine and plasma samples in both positive and negative ion modes. These relevant variables, which were significantly changed in response to these effects, were selected based on the calculation of the significance of the variables (P value < 0.05). In addition to the calculation of the significance of the variables (paired t-test, p-value,), all significant features were evaluated for significance using significance of variables in prediction (VIP) statistics. The significance of the VIP projection coefficient variables reflects the contribution of hidden variables to the dependent model in relation to other variables through assessment of the relative importance of each X variable in relation to each X variable in the predictive model. Metabolites with a VIP value greater than 1 (VIP > 1) this mean it contribute the most to in the predictive model **[17]**.

Significantly differences in thirty-five urinary metabolites and thirty-three plasma metabolites were identified by two-tailed t-test with threshold 0.05. Metabolites with notable differences in their profiles in comparing of urine samples and plasma samples, both samples separately, pre- and post-intake beetroot juice are summarised in table 1 and 2, respectively. Most metabolic pathways including amino acid, peptide, Co-factors and vitamins were significantly affected. Thus, these significant metabolites were projected as key features in section of data modelling for this study.

Table 1: The relevant important metabolites with high impact in plasma samples at pre and post-intake beetroot juice based on the critical threshold for a regarding a P-value as being significant is 0.05 and VIP≥1.0. N= negative ion and P.

N	Ion Mode	m/z	RT	Molecular formula	Name	CV Qc %	VIP	P-Value C1-C2	Ratio C2/C1	P-Value C1-C4	Ratio C4/C1	P-Value C1-C8	Ratio C8/C1
1	P728	384.274	5.67	C21H37NO5	3-Hydroxy-5, 8-tetradecadiencarnitine	13.41	1.63	0.0113	0.556	0.0162	0.585	0.0191	0.614
2	P674	368.279	5.07	C21H37NO4	3, 5-Tetradecadiencarnitine	10.91	2.01	0.0011	0.415	0.0029	0.468	0.0007	0.391
3	P365	386.290	5.54	C21H39NO5	3-Hydroxy-cis-5-tetradecenoylcarnitine	8.75	1.91	0.0066	0.572	0.0036	0.538	0.0014	0.484
4	P154	342.264	5.27	C19H35NO4	trans-2-Dodecenoylcarnitine	18.48	1.91	0.0028	0.417	0.0037	0.439	0.0026	0.394
5	P593	312.217	5.75	C17H29NO4	2-trans,4-cis-Decadienoylcarnitine	16.24	1.77	0.0037	0.568	0.0049	0.598	0.0028	0.547
6	P547	314.232	5.69	C17H31NO4	9-Decenoylcarnitine	6.07	1.95	0.0003	0.501	0.0026	0.569	0.0006	0.501
7	P1020	330.227	6.78	C17H31NO5	6-Keto-decanoylcarnitine	10.87	1.73	0.0470	0.758	0.0173	0.702	0.0015	0.624
8	P794	258.170	7.93	C13H23NO4	2-Hexenoylcarnitine	10.91	1.39	0.0204	0.773	0.0209	0.785	0.0496	0.843
9	N413	232.028	13.75	C8H11NO5S	Dopamine 3-O-sulfate	11.11	2.01	0.0000	42.116	0.0000	51.588	0.0000	11.821
10	P332	330.264	5.39	C18H35NO4	4-8dimethylnonanoylcarnitine	10.62	1.87	0.0006	0.515	0.0019	0.552	0.0018	0.529
11	P826	372.311	4.97	C21H41NO4	Tetradecanoylcarnitine	14.49	1.61	0.0227	0.758	0.0038	0.676	0.0174	0.737
12	N335	209.067	12.46	C7H14O7	Sedoheptulose	13.64	2.00	0.0000	3.844	0.0000	7.368	0.0000	4.836
13	N63	119.035	10.64	C4H8O4	D-Erythrose	19.03	1.22	0.0101	0.839	0.0327	0.872	0.0393	0.901
14	P138	370.295	5.01	C21H39NO4	cis-5-Tetradecenoylcarnitine	7.05	1.82	0.0259	0.482	0.0135	0.425	0.0076	0.379
15	N584	225.186	5.02	C14H26O2	(9Z)-Tetradecenoic acid	17.52	2.06	0.0032	0.548	0.0011	0.490	0.0037	0.562
16	P515	195.076	5.64	C9H10N2O3	4-Aminohippuricacid	13.12	2.07	0.0000	95.650	0.0000	60.206	0.0007	12.553
17	N485	144.030	8.87	C5H7NO4	2-Oxoglutaramate	20.05	1.42	0.0036	1.223	0.0000	1.413	0.0339	1.212
18	N374	178.072	11.92	C6H13NO5	D-Glucosamine	19.19	1.89	0.0000	1.930	0.0000	1.560	0.0249	1.233
19	P1	118.086	11.82	C5H11NO2	L-Valine	14.62	1.97	0.0000	1.839	0.0000	1.621	0.0364	1.247
20	P140	288.217	6.01	C15H29NO4	L-Octanoylcarnitine	13.83	1.88	0.0033	0.449	0.0043	0.472	0.0042	0.445
21	P595	260.186	7.38	C13H25NO4	[FA (6:0)] O-hexanoyl-R-carnitine	9.63	1.92	0.0005	0.534	0.0004	0.513	0.0042	0.619

22	N447	361.202	5.21	C21H30O5	[ST trihydroxy(2:0)] 11beta,17,21-trihydroxypregn-4-ene-3,20-dione	8.91	1.78	0.0122	0.781	0.0259	0.828	0.0001	0.642
23	N162	331.192	5.20	C20H28O4	Gibberellin A12	8.91	1.62	0.0215	0.739	0.0264	0.782	0.0003	0.622
24	N246	169.123	5.68	C10H18O2	[PR] Limonene-1,2-diol	13.38	1.97	0.0002	0.619	0.0011	0.664	0.0006	0.657
25	P538	316.248	5.59	C17H33NO4	[FA (10:0)] O-decanoyl-R-carnitine	5.65	1.86	0.0049	0.394	0.0092	0.450	0.0050	0.386
26	N416	162.940	9.99	C4H5O2Br	[FA (4:0)] 2-bromo-2-butenoic acid	13.85	2.12	0.0000	11.985	0.0000	11.399	0.0000	7.899
27	N443	274.039	5.33	C10H13NO6S	L-Tyrosine methyl ester 4-sulfate	9.71	1.94	0.0000	9.870	0.0000	12.786	0.0000	10.386
28	P187	344.279	5.21	C19H37NO4	1,2-dioctanoyl-1-amino-2,3-propanediol	10.89	2.09	0.0015	0.500	0.0006	0.455	0.0004	0.435
29	N481	233.012	6.37	C8H10SO6	trihydroxy phenylethanol sulfate isomer	14.51	1.76	0.0000	2.984	0.0001	3.113	0.0164	2.147
30	N807	230.997	6.10	C8H8SO6	hydroxyphenyl acetic acid sulfate or isomer	8.62	1.60	0.0004	2.998	0.0001	6.406	0.0000	3.744
31	P301	177.075	10.32	C7H12O5	(2S)-2-Isopropylmalate	19.72	1.38	0.0393	0.884	0.0010	0.854	0.0458	0.904
32	P817	205.043	6.29	C12H9OCl	4-Chloro-4'-biphenylol	12.98	1.57	0.0001	20.022	0.0027	25.142	0.0020	2.518
33	P1071	118.121	11.87	C6H15NO	2-Methylcholine	9.32	1.92	0.0000	1.826	0.0000	1.617	0.0433	1.232

Table 2: The relevant important metabolites with high impact in Urine samples at pre and post-intake beetroot juice based on the critical threshold for a regarding a P-value as being significant is 0.05 and VIP≥1.0. N= negative ion and P.

No	Ion Mode	M/Z	RT	Molecular formula	Name	CV-Qc %	VIP	P-Value UR1-UR2	Ratio UR2/R1	P-Value UR1-UR3	Ratio UR3/R1	P-Value UR1-UR4	Ratio UR4/R1
1	P9	180.102	11.07	C10H13NO2	(-)-Salsolinol	19.3	1.67	0.0000	5.280	0.0006	19.284	0.0000	10.859
2	P73	152.071	7.66	C8H9NO2	(Z)-4-Hydroxyphenylacetaldehyde-oxime	4.1	1.41	0.0004	24.385	0.0000	24.848	0.0001	6.861
3	P75	152.071	9.30	C8H9NO2	(Z)-4-Hydroxyphenylacetaldehyde-oxime	16.5	1.36	0.0014	1.701	0.0000	1.774	0.0254	1.228
4	P426	156.066	8.18	C7H9NO3	2-amino-5-methyl-muconate semialdehyde	8.6	1.69	0.0000	4.346	0.0000	17.785	0.0003	6.378
5	P553	154.050	7.05	C7H7NO3	3-Hydroxyanthranilate	14.7	1.62	0.0015	2.592	0.0000	7.964	0.0024	2.722

6	P643	195.076	5.65	C9H10N2O3	4-Aminohippuricacid	8.6	1.79	0.0000	53.568	0.0000	83.148	0.0000	30.230
7	P673	177.066	5.63	C9H8N2O2	4-Hydroxyaminoquinoline N-oxide	11.6	1.78	0.0000	46.459	0.0000	73.571	0.0000	25.241
8	P678	110.060	5.58	C6H7NO	4-Hydroxyaniline	17.5	1.76	0.0000	7.288	0.0000	11.660	0.0000	4.223
9	P762	193.097	7.57	C10H12N2O2	5'-Hydroxycotinine	17.0	1.30	0.0011	7.383	0.0000	7.155	0.0003	2.537
10	P773	221.092	6.46	C11H12N2O3	5-Hydroxy-L-tryptophan	12.0	1.46	0.0054	1.959	0.0000	2.409	0.0064	1.656
11	P12227	224.092	12.11	C11H13NO4	Bendiocarb	18.3	1.91	0.0000	4.207	0.0000	19.160	0.0001	18.740
12	P12388	232.097	7.43	C13H13NO3	Benzoyldehydro-2,3-dihydroxy-benzone	3.9	1.73	0.0003	2.751	0.0000	5.308	0.0123	2.276
13	P13955	222.087	8.86	C8H16NO4P	D,L-cyclohexanephosphinothricin	18.4	1.60	0.0000	10.818	0.0001	15.481	0.0001	4.476
14	P18922	133.076	5.62	C8H8N2	Indoleamine	10.5	1.78	0.0000	73.757	0.0000	118.601	0.0000	41.026
15	P18933	133.076	8.30	C8H8N2	Indoleamine	18.2	1.52	0.0001	6.434	0.0002	12.514	0.0300	2.885
16	P19611	289.085	13.51	C19H12O3	lachnanthocarpone	8.4	1.74	0.0000	10.457	0.0000	19.286	0.0000	8.374
17	P21000	209.092	8.57	C10H12N2O3	L-Kynurenine	10.3	1.38	0.0001	2.806	0.0001	3.163	0.0019	1.872
18	P21755	120.066	15.79	C4H9NO3	L-Threonine	19.5	1.48	0.0432	1.315	0.0003	1.704	0.0400	1.232
19	P26368	124.040	8.02	C6H5NO2	Nicotinate	15.3	1.36	0.0022	51.819	0.0000	56.021	0.0005	10.674
20	P26433	181.061	8.89	C8H8N2O3	Nicotinurate	15.6	1.64	0.0001	13.098	0.0000	24.316	0.0001	4.654
21	P27599	158.081	8.64	C7H11NO3	Paramethadione	17.3	1.63	0.0245	1.354	0.0000	2.933	0.0007	2.009
22	P29990	168.066	8.65	C8H9NO3	Pyridoxal	11.1	1.73	0.0000	6.075	0.0000	8.016	0.0000	3.568
23	P30422	165.066	8.07	C8H8N2O2	Ricinine	3.5	1.50	0.0003	4.251	0.0000	4.651	0.0000	1.961
24	P33044	144.066	17.57	C6H9NO3	Vinylacetylglycine	17.0	1.71	0.0105	1.434	0.0000	2.243	0.0007	1.820
25	N42	237.052	8.15	C10H10N2O5	(3-nitrobenzoyl)alanine	13.3	1.29	0.0007	6.937	0.0041	87.611	0.0188	15.294
26	N829	193.062	5.65	C9H10N2O3	4-Aminohippuricacid	9.2	1.80	0.0000	41.158	0.0000	62.632	0.0000	24.633
27	N10008	179.038	10.33	C6H12O4S	5-Methylthio-D-ribose	11.4	1.33	0.0183	0.647	0.0027	0.560	0.0039	0.562
28	N11877	487.178	15.78	C19H32N6O5S2	Ala-Met-Met-His	12.3	1.63	0.0000	4.109	0.0001	7.423	0.0040	2.272

29	N1222	175.047	10.21	C4H8N4O4	Allantoate	16.1	1.71	0.0002	1.862	0.0000	2.092	0.0002	2.135
30	N1252	136.040	14.94	C7H7NO2	Anthranilate	3.4	1.77	0.0027	1.747	0.0000	2.859	0.0002	1.930
31	N1254	136.040	8.32	C7H7NO2	Anthranilate	12.7	1.62	0.0000	3.292	0.0000	5.451	0.0068	1.972
32	N21733	252.129	7.85	C11H19N5S	Irgarol 1051	18.2	1.68	0.0314	0.280	0.0083	0.120	0.0286	2.221
33	N27888	324.073	13.75	C14H15NO8	Pancratistatin	9.0	1.92	0.0003	4.184	0.0000	18.026	0.0006	34.281
34	N27997	156.067	8.65	C7H11NO3	Paramethadione	11.0	1.75	0.0117	1.366	0.0001	2.780	0.0001	2.136
35	N29551	149.072	5.67	C8H10N2O	p-nitroso-N,N-dimethylaniline	8.9	1.78	0.0000	30.914	0.0000	49.076	0.0000	17.550

Data Modelling by Unsupervised and Supervised Models:

A multivariate data modelling analysis was performed using two general modelling strategies to identify patterns of metabolic differences. Unsupervised PCA models and supervised OPLS-DA models were created for each set of biofluid samples to compare beet juice metabolites

Unsupervised Data Modelling:

The data obtained from the analysis of plasma and urine samples on a ZICpHILIC column were first modeled using a principal component analysis, unsupervised model, to reduce the individual components and obtain some form of linear summation by determining the exogenous variables. Principal component analysis (PCA) is defined as ″a multivariate technique that analyzes a data table in which observations are described by several inter-correlated quantitative dependent variables. Its goal is to extract the important information from the table, to represent it as a set of new orthogonal variables called principal components, and to display the pattern of similarity of the observations and of the variables as points in maps″ **[18]**.

In addition to the overview of compounds separation by model, the main diagnostic tools that provide a summary of the suitability and goodness of the PCA model are Variance Explained - R2X (Cum) and Variance Predicted-Q2 (Cum) **[19 - 21]**. Table 3 summarizes the R2 and Q2 values as diagnostic tools of model descriptions for changes in both samples, plasma and urine, to indicate to model goodness and fit dependent on significant discrepancies values.

Table 3: summarising the R2 and Q2 values in both plasma and urine samples.

Model / Specimen	PCA		
	Group	R2X(Cum)	Q2(Cum)
Plasma	Pre- vs post-	0.527	0.571
Urine	Pre- vs post-	0.813	0.692

Figure 2 shows the split PCA of the four groups in beetroot juice taking study after normalization of individual metabolites from plasma samples collected at four time points per subjects. While, figure 3 shows same analysis of PCA for urine samples at same four time points as in plasma samples.

Based on the overview of the observations in Figures 1 and 2, we can see that PCA was not clearly segregated due to the unobserved differences in metabolites profile between plasma samples unlike urine samples that were clear separation by PCA as a result to observed differences in metabolites profile. Also, the R2X (0.527 and 0.813) and Q2 (0.571 and 0.692) values for plasma and urine, respectively, presented in table 3 match with the overview of the observations in figures 1 and 2 and confirm that the PCA separation for urine samples more clear and fit than plasma samples.

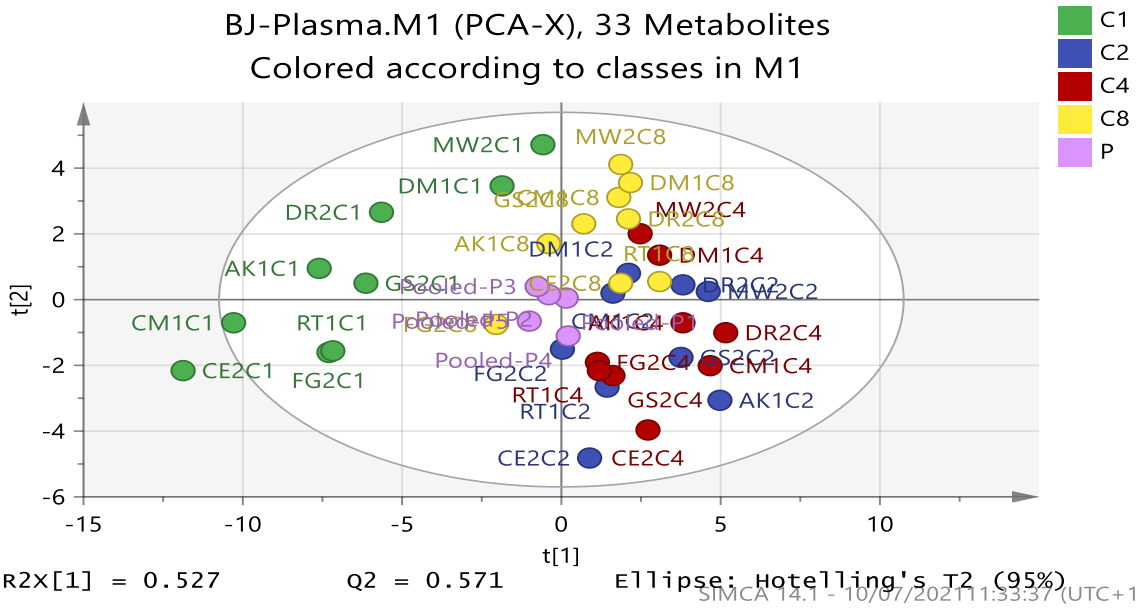


Figure 2: The split PCA of the four plasma sample groups in beetroot juice taking study.

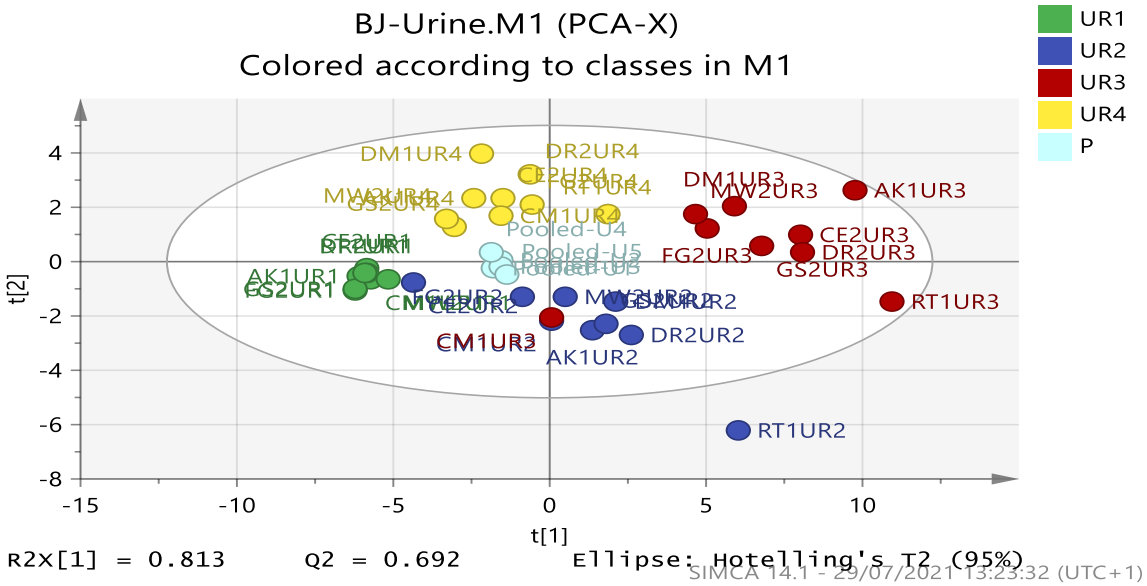


Figure 3: The split PCA of the four urine sample groups in beetroot juice taking study.

Observation maps (point plots applied by PCA) are commonly used to compress single components and unreal variables, aggregate linear input data, and exploit data collinearity **[22]**. Once the grading scheme has been achieved to fit the data and dimensional issues have been eliminated, the OPLS-DA classification model is used to divide the sample into each group to understand differences between groups. Therefore, supervised data modelling will be required.

Supervised Data Modelling of Urine Samples:

If identification of significant split between study groups was not achieved, Reverting to a supervised OPLS-DA model may be necessary to determinate statistical effects. OPLS-DA can identify splits that PCA cannot, but the statistical significance of the splits must be tested before conclusions can be drawn from the results **[23]**. The score plot of OPLS-DA model was performed to comparison between pre-taken samples and post-taken at three time points for urine samples. This comparison goals to evaluate the highly metabolomic changes between study groups to establish fit OPLS-DA model. According to modelling results and clear separation by PCA, the comparison between urine sample groups is appropriate selection to choose the best time point can be used in OPLS-DA analysis. Three OPLS-DA models of the data were built for this comparison in order to get a clearer picture of the differences between the first sample (Ur1) and second (Ur2), third (Ur3) and fourth (Ur4) urine samples as show on figures 4, 5 and 6. The outlier variables were observed in figure 4 (CE2UR2, red color) and figure 6 (RT1UR4, red color) which reduce the goodness of model fitting.

The overview of OPLS-DA model and model quality characterized by R2X, Q2 and CV-ANOVA (P-Value) values which are detailed in table 4 illustrate that the comparison between Ur1 and Ur3 samples is best time points to build multivariate models for our study approaches.

Table 4: summarising the R2, Q2 and P-values for OPLS-DA in urine samples.

OPLS-DA for Urine samples			
Time point	R2X(Cum)	Q2(Cum)	P-Value
Ur1 vs Ur2	0.765	0.743	0.00157942
Ur1 vs Ur3	0.837	0.839	0.000106282
Ur1 vs Ur4	0.797	0.787	0.000480965

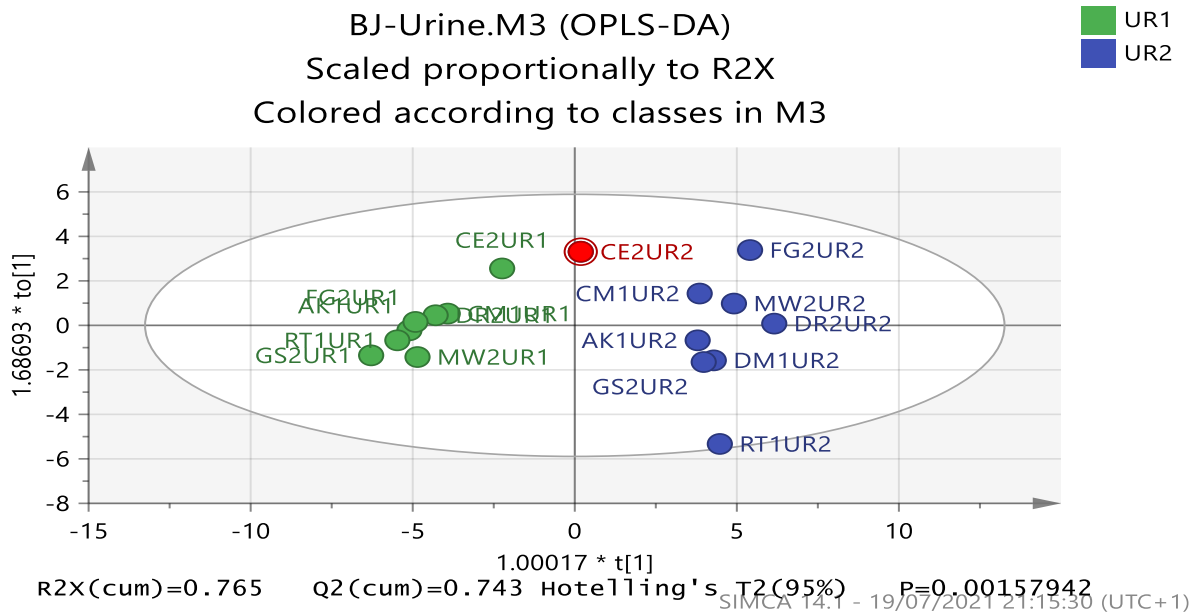


Figure 4: Overview of the observations of OPLS-DA model for separation between two time points (Ur1 and Ur2) for urine samples from healthy subject. R2X (cum) 0.765, Q2 (cum) 0.743 and P-value= 0.00157.

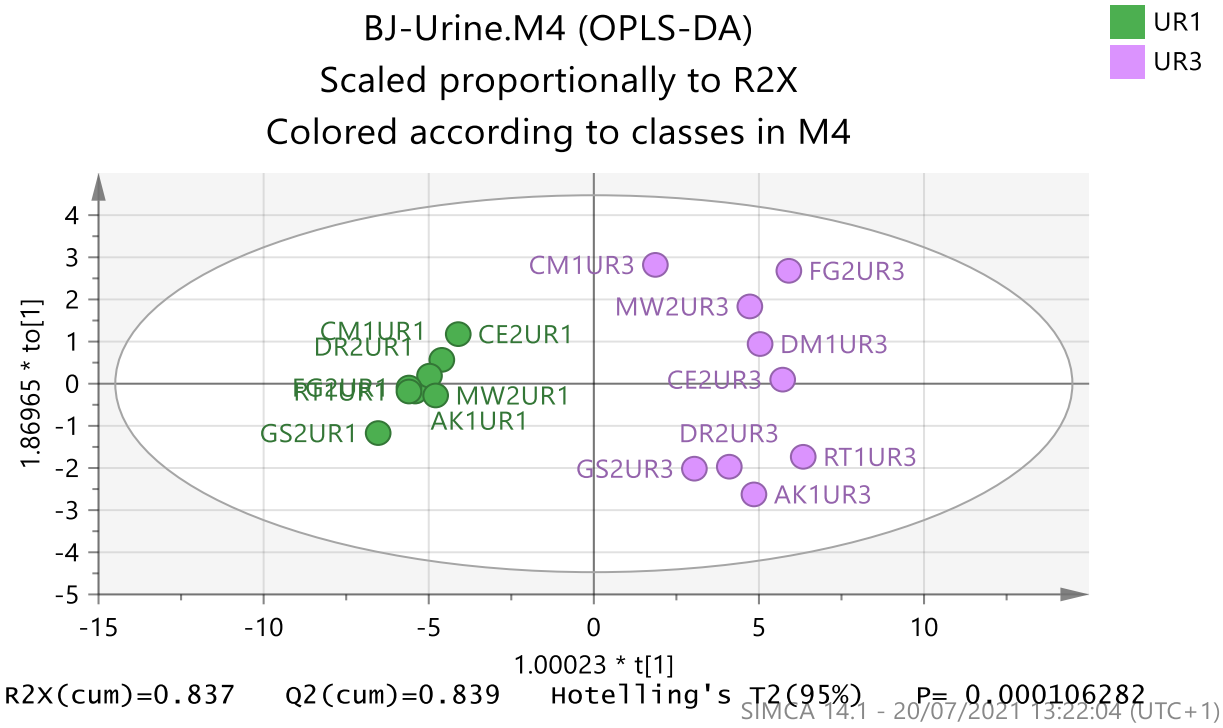


Figure 5: Overview of the observations of OPLS-DA model for separation between two time points (Ur1 and Ur3) for urine samples from healthy subject. R2X (cum) 0.837, Q2 (cum) 0.839 and P-value= 0.000106

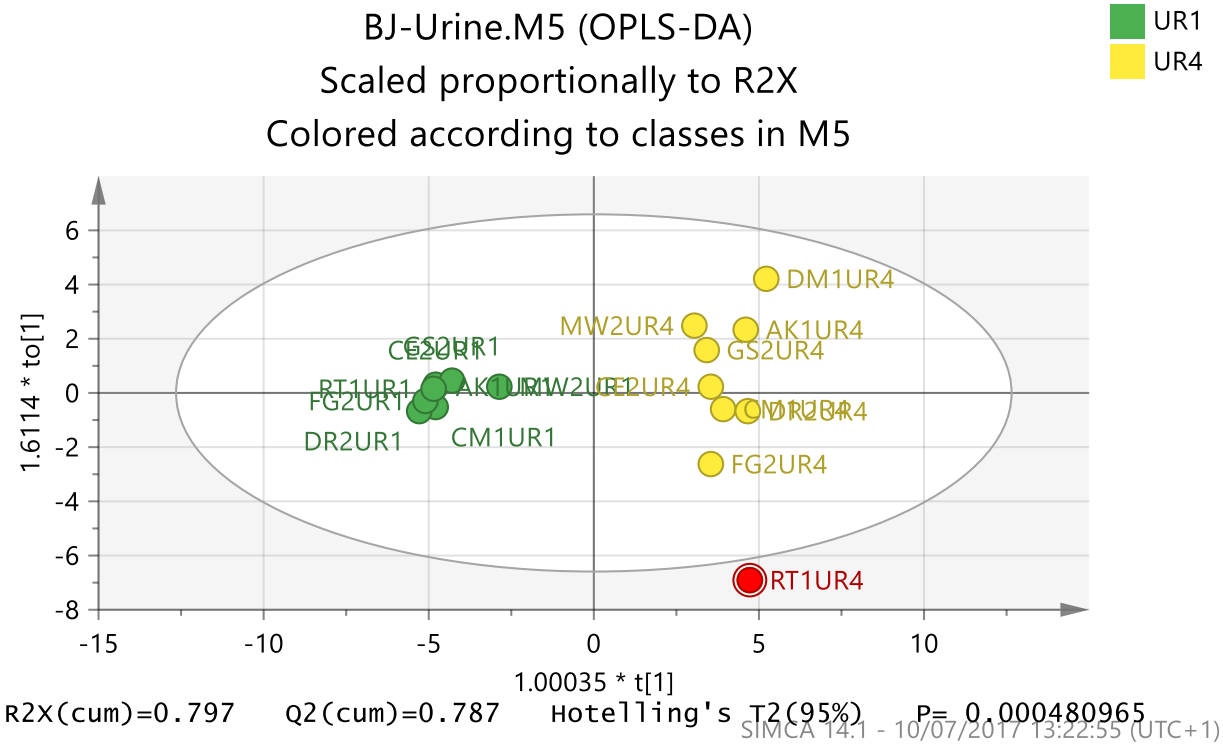


Figure 6: Overview of the observations of OPLS-DA model for separation between two time points (Ur1 and Ur4) for urine samples from healthy subject. R2X (cum) 0.797, Q2 (cum) 0.787 and P-value= 0.00048.

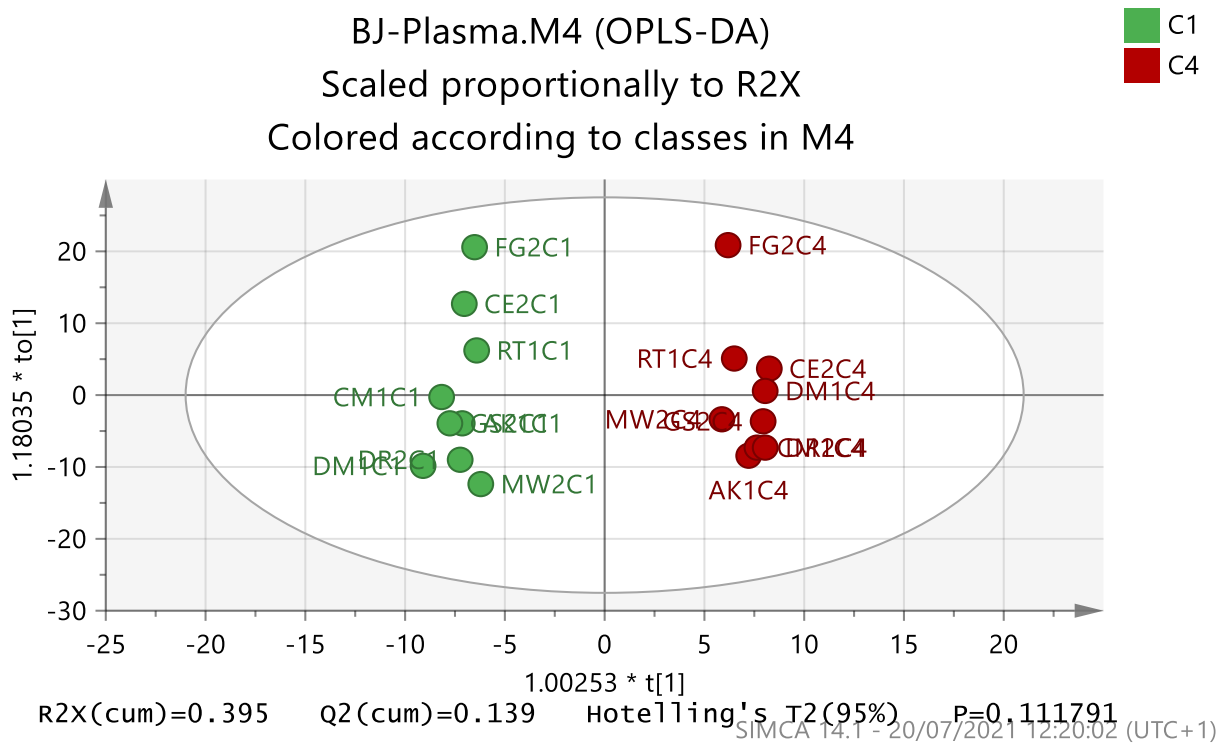


Figure 7: Overview of the observations of OPLS-DA model for separation between two times points (C1 and C4) for plasms samples from healthy subject. R2X (cum) 0.395, Q2 (cum) 0.139 and P-value= 0.111791.

Supervisor Data Modelling of Urine vs Plasma Samples:

The score plot of OPLS-DA model was carried out in this step from current study to comparison between baseline samples and post-taken samples at one time point, after 120min, for plasma (C1 vs C4), figure 7, and urine (Ur1 vs Ur3) samples, figure 5, to assess the goodness of OPLS-DA data modelling.

Due to internal validation methods such as interpretations of variance and expected values as well as overview of model are not sufficient to assess model quality and fitting. In untargeted metabolomics studies, the external validation of data modeling is necessary to obtain reliable results **[23]**. Two external validation methods, including cross-validation ANOVA (CV-ANOVA) and the permutation test can be performed for OPLS-DA models generated with SIMCA-P. CV-ANOVA presents the returned P-value which indicates the statistical significance of the test with threshold 0.005. While, a reference distribution for R2/Q2 values is provided by permutation test **[24]**. The model validity criterion according to permutation test is achieved under two conditions. First, all green values in R2 must be less than the starting point on the right. Then all the Q2 values (blue) on the left should be less than the right side points (original point), or the Q2 points regression line intersects the left vertical axis to be under zero **[25]**.

CV-ANOVA with p-value, first external validation methods, was calculated for two OPLS-DA models built for obtained data from plasma samples (C1 vs C4) and urine samples (Ur1 vs Ur3). These values confirm that model which built for plasma samples was not fitting and nonsignificant with P-value at 0.111791. Contrastingly, the returned P-value of OPLS-DA model built for urine samples was 0.000106282 which indicates that this model is high significantly fitting.

Through application the second external validation methods, the permutation test gave the same outcome and provided the evidence on that OPLS-DA models are not fitting in plasma sample (figure 8) according to permutation test conditions. It is observed, all green values in R2 are in same levels with, not less than, the starting point on the right. Also, many of the blue Q2 values on the left are high than the original points on the right. This observations indicate to weakness of OPLS-DA model for plasma samples and it is not fitted to be accepted.

Two conditions of permutation test were achieved in the model validity criterion for the comparison between urine samples. This test gave the reliable assessment for the goodness of fit (R2 and Q2) of the original model at this comparison between Ur1 vs Ur3 as shown on figure 9.

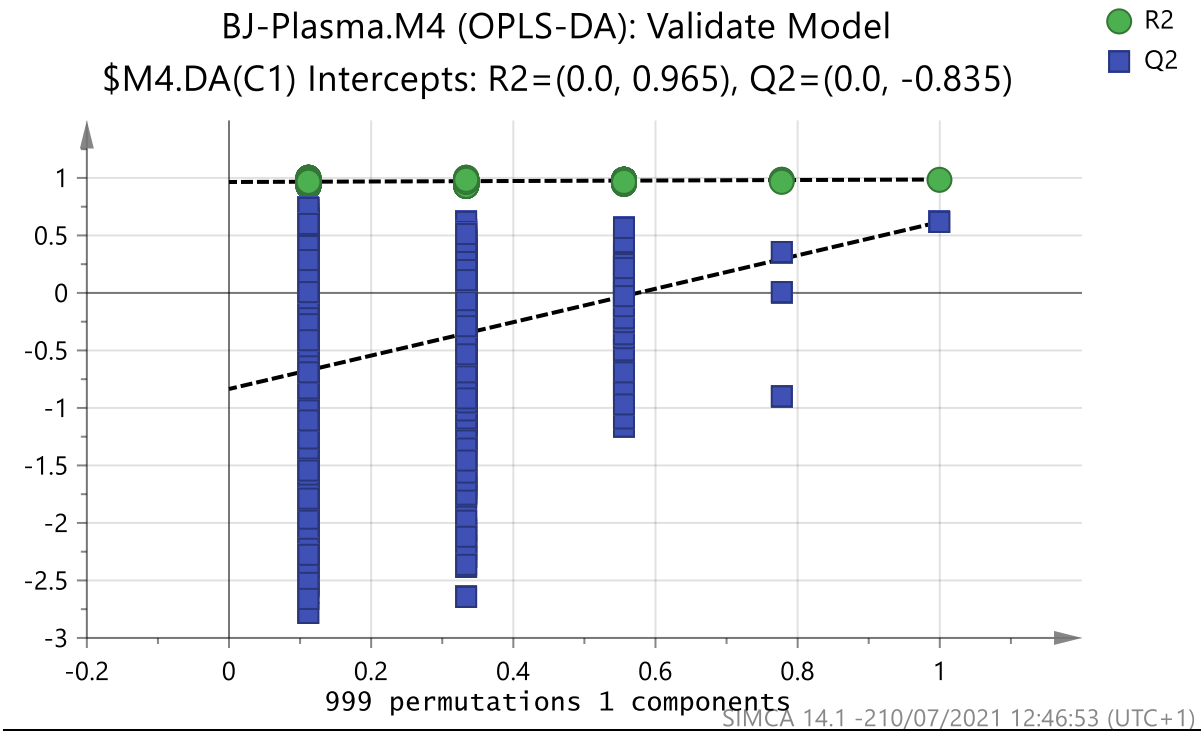


Figure 8: Cross validation of OPLSDA model for the classification of pre-beetroot taking group (C1-C4) in plasma samples by the Permutations test.

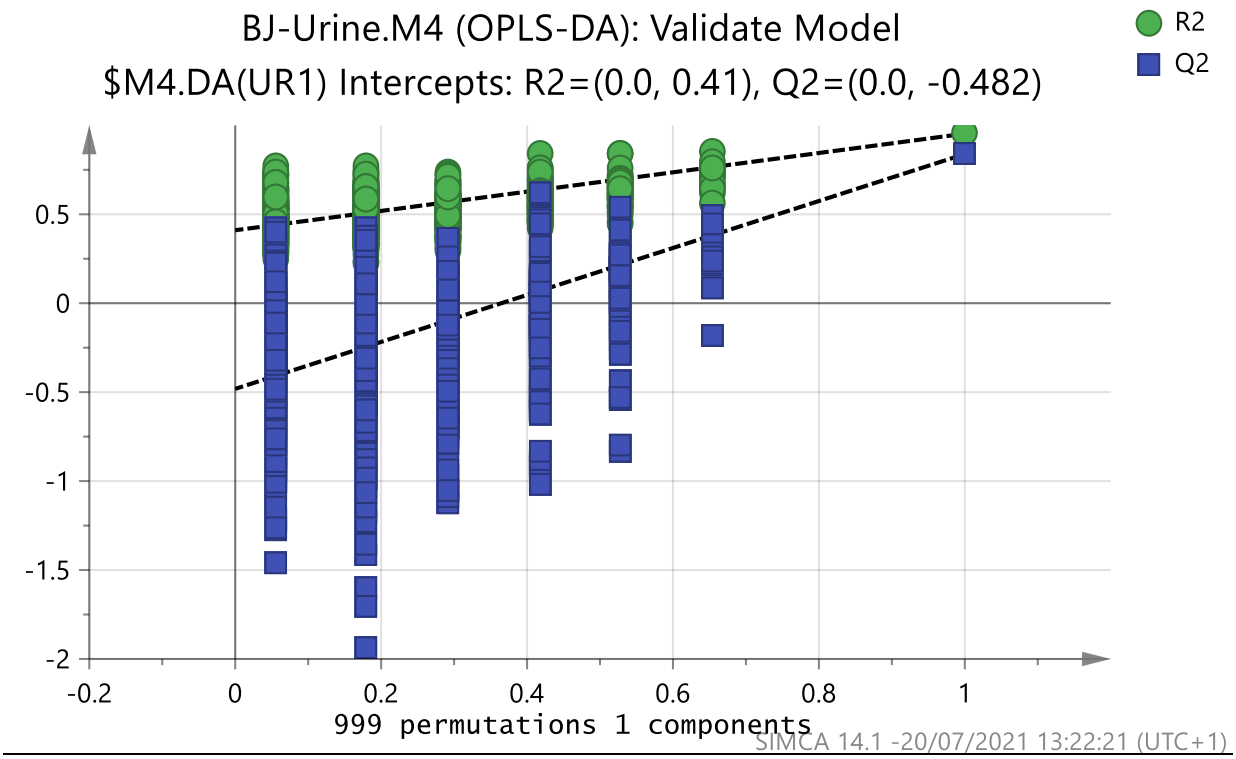


Figure 9: Cross validation of OPLSDA model for the classification of pre-and post-beetroot taking group (Ur1-Ur3) in urine samples by the Permutations test.

Conclusions:

An LC/MS-based current pilot study which designed to assess potential beetroot juice effect on human metabolome showed that metabolic changes were more pronounced in urine samples than in plasma samples. Study outcomes that reflected acute metabolic patterns in urine compared with plasma suggest that metabolic screening of urine samples before and after nutrition with beetroot juice is the best possible tool for predicting the metabolic characterization of this nutrition effect. Reasonable approach.

The use of PCA and OPLS-DA in study data modelling shows a clear separation of urinary metabolites and provides a reasonable approach to distinguish between pre- and post-intake beetroot juice groups, although separation would not be possible in plasma groups by same approach.

Due to this study designed to investigate the potential specimen and appropriate approach which can be use as the prospective tool to reveal the metabolic effect of a diet pre- and post-intake beetroot juice, it focused on one group of ages and gender. Therefore, as a recommendation, an influence of age and gender on the human plasma and urine samples associated to nutrition (beetroot juice) is recommended in future.

Also, throughout the comparison study between changed metabolites as a result of beetroot effects taking, we noted there were several metabolic pathways are affected after beetroot juice digestion. These metabolic pathways included amino acids, lipids, peptides, xenobiotics, carbohydrates, nucleotides, cofactors and vitamins that need to be studied more depth in future. Thus, this future work may be very important in order to study the impact of beetroot on human metabolomic profiling and explore potential predictive markers for these effects.

References:

1. Ohlhorst, S. D., Russell, R., Bier, D., Klurfeld, D. M., Li, Z., Mein, J. R., ... & Konopka, E. (2013). Nutrition research to affect food and a healthy life span. The Journal of nutrition, 143(8), 1349-1354.
2. Santamaria, P. (2006). Nitrate in vegetables: toxicity, content, intake and EC regulation. Journal of the Science of Food and Agriculture, 86(1), 10-17.
3. Joshipura, K. J., Hu, F. B., Manson, J. E., Stampfer, M. J., Rimm, E. B., Speizer, F. E., ... & Willett, W. C. (2001). The effect of fruit and vegetable intake on risk for coronary heart disease. Annals of internal medicine, 134(12), 1106-1114.
4. Joshipura, K. J., Ascherio, A., Manson, J. E., Stampfer, M. J., Rimm, E. B., Speizer, F. E., ... & Willett, W. C. (1999). Fruit and vegetable intake in relation to risk of ischemic stroke. Jama, 282(13), 1233-1239.
5. Coles, L. T., & Clifton, P. M. (2012). Effect of beetroot juice on lowering blood pressure in free-living, disease-free adults: a randomized, placebo-controlled trial. Nutrition journal, 11(1), 1-6
6. Aliahmadi, M., Amiri, F., Bahrami, L. S., Hosseini, A. F., Abiri, B., & Vafa, M. (2021). Effects of raw red beetroot consumption on metabolic markers and cognitive function in type 2 diabetes patients. Journal of Diabetes & Metabolic Disorders, 20(1), 673-682.
7. Brennan, L. (2016). Metabolomics in nutrition research–a powerful window into nutritional metabolism. Essays in Biochemistry, 60(5), 451-458.
8. Orešič, M. (2009). Metabolomics, a novel tool for studies of nutrition, metabolism and lipid dysfunction. Nutrition, Metabolism and Cardiovascular Diseases, 19(11), 816-824
9. Gibbons, H., & Brennan, L. (2017). Metabolomics as a tool in the identification of dietary biomarkers. Proceedings of the Nutrition Society, 76(1), 42-53
10. Issaq, H.J., et al., Analytical and statistical approaches to metabolomics research. Journal of separation science, 2009. 32(13): p. 2183-2199.
11. Bruce, S.J., et al., Evaluation of a protocol for metabolic profiling studies on human blood plasma by combined ultra-performance liquid chromatography/mass spectrometry: From extraction to data analysis. Analytical Biochemistry, 2008. 372(2): p. 237-249.
12. Zhang, T., et al., Application of Holistic Liquid Chromatography-High Resolution Mass Spectrometry Based Urinary Metabolomics for Prostate Cancer Detection and Biomarker Discovery. PLoS One, 2013. 8(6): p. e65880.
13. Zhang, R., et al., Metabolomic Profiling of Post-Mortem Brain Reveals Changes in Amino Acid and Glucose Metabolism in Mental Illness Compared with Controls. Computational and structural biotechnology journal, 2016. 14: p. 106-116.
14. Pluskal, T., et al., MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics, 2010. 11: p. 395.
15. Muhsen Ali, A., et al., Metabolomic Profiling of Submaximal Exercise at a Standardised Relative Intensity in Healthy Adults. Metabolites, 2016. 6(1).
16. Alonso, A., S. Marsal, and A. Julià, Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. Frontiers in Bioengineering and Biotechnology, 2015. 3: p. 23.
17. Chong, I.-G. and C.-H. Jun, Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems, 2005. 78(1–2): p. 103-112.
18. Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley interdisciplinary reviews: computational statistics, 2(4), 433-459.
19. Wheelock, Å.M. and C.E. Wheelock, Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. Molecular bioSystems, 2013. 9(11): p. 2589-2596.

20. Westerhuis, J.A., et al., Assessment of PLSDA cross validation. Metabolomics, 2008. 4(1): p. 81-89.
21. Triba, M.N., et al., PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. Molecular BioSystems, 2015. 11(1): p. 13-19.
22. Liland, K.H., Multivariate methods in metabolomics – from pre-processing to dimension reduction and statistical analysis. TrAC Trends in Analytical Chemistry, 2011. 30(6): p. 827-841.
23. K Trivedi, D., The Application of SIMCA P+ in Shotgun Metabolomics Analysis of ZIC®HILIC-MS Spectra of Human Urine - Experience with the Shimadzu IT-T of and Profiling Solutions Data Extraction Software. Journal of Chromatography & Separation Techniques, 2012. 03(06).
24. Worley, B. and R. Powers, Multivariate Analysis in Metabolomics. Curr Metabolomics, 2013. 1(1): p. 92-107.
25. Eriksson, L., J. Trygg, and S. Wold, CV-ANOVA for significance testing of PLS and OPLS® models. Journal of Chemometrics, 2008. 22(11-12): p. 594-600.